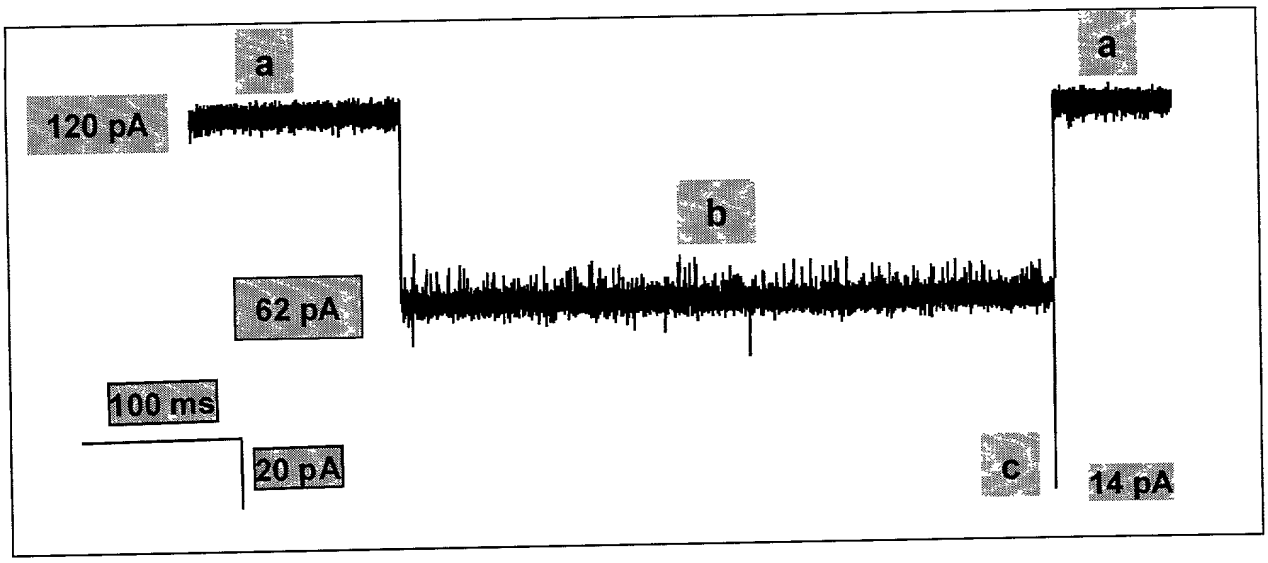


Figure 1

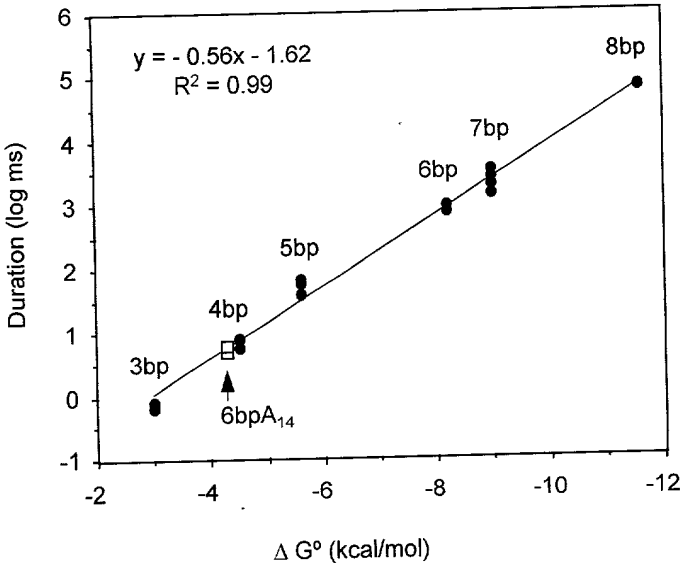

y = - 0.56x - 1.62
R^2 = 0.99
8bp
7bp
6bp
5bp
4bp
3bp
6bpA14
Duration (log ms)
Δ G° (kcal/mol)
6
5
4
3
2
1
0
-1
-2
-4
-6
-8
-10
-12

Figure 2

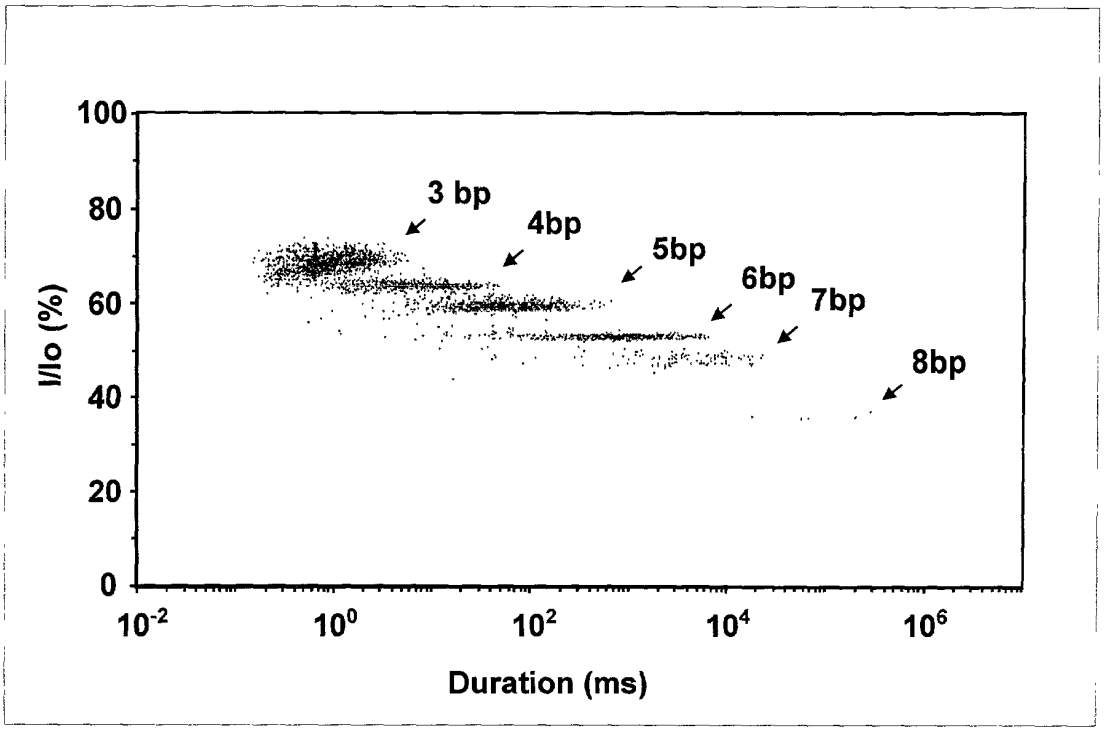

100
80
60
40
20
0
I/Io (%)
3 bp
4bp
5bp
6bp
7bp
8bp
10^{-2}
10^{0}
10^{2}
10^{4}
10^{6}
Duration (ms)

Figure 3a

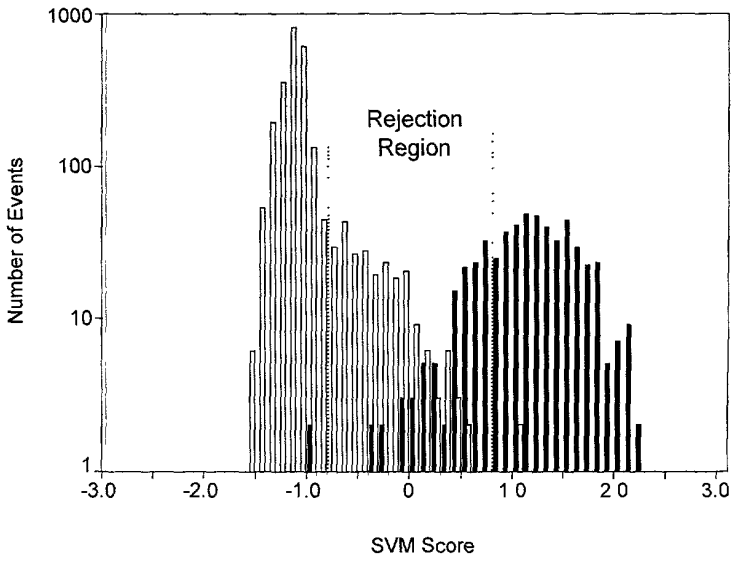

1000
100
10
1
Number of Events
Rejection
Region
-3.0
-2.0
-1.0
0
1 0
2 0
3.0
SVM Score

Figure 3b

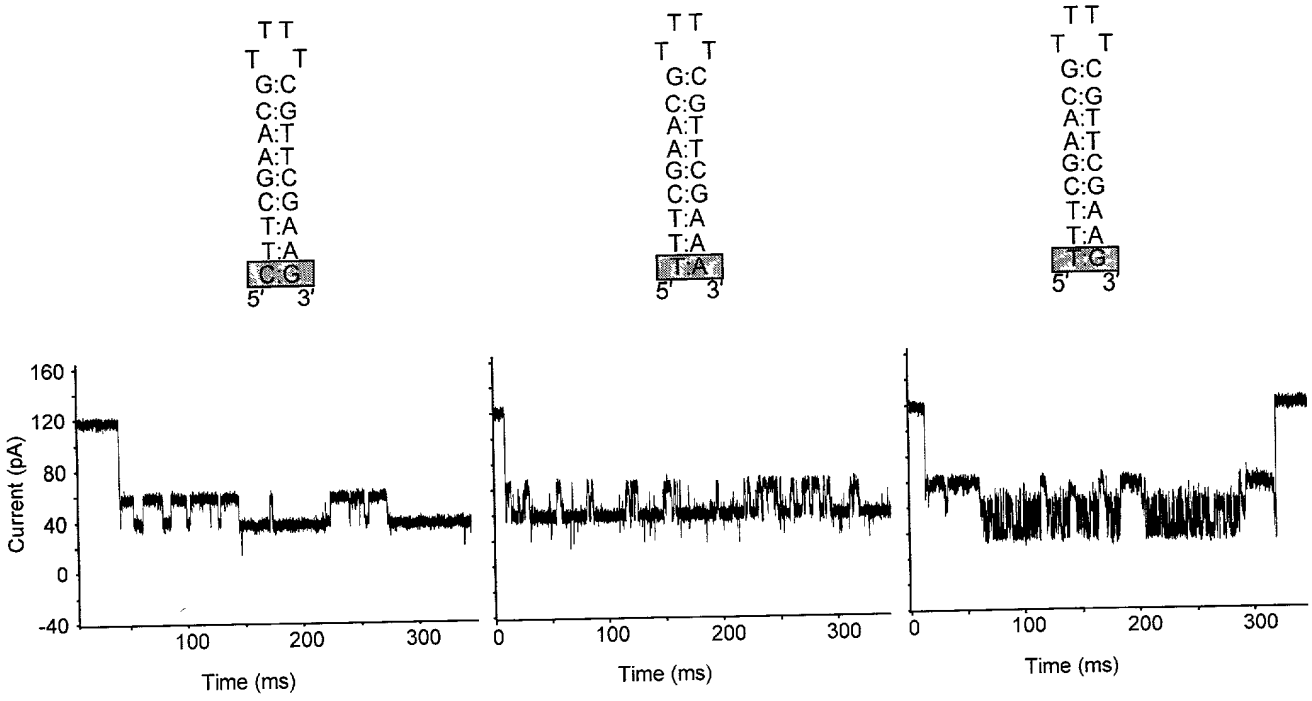
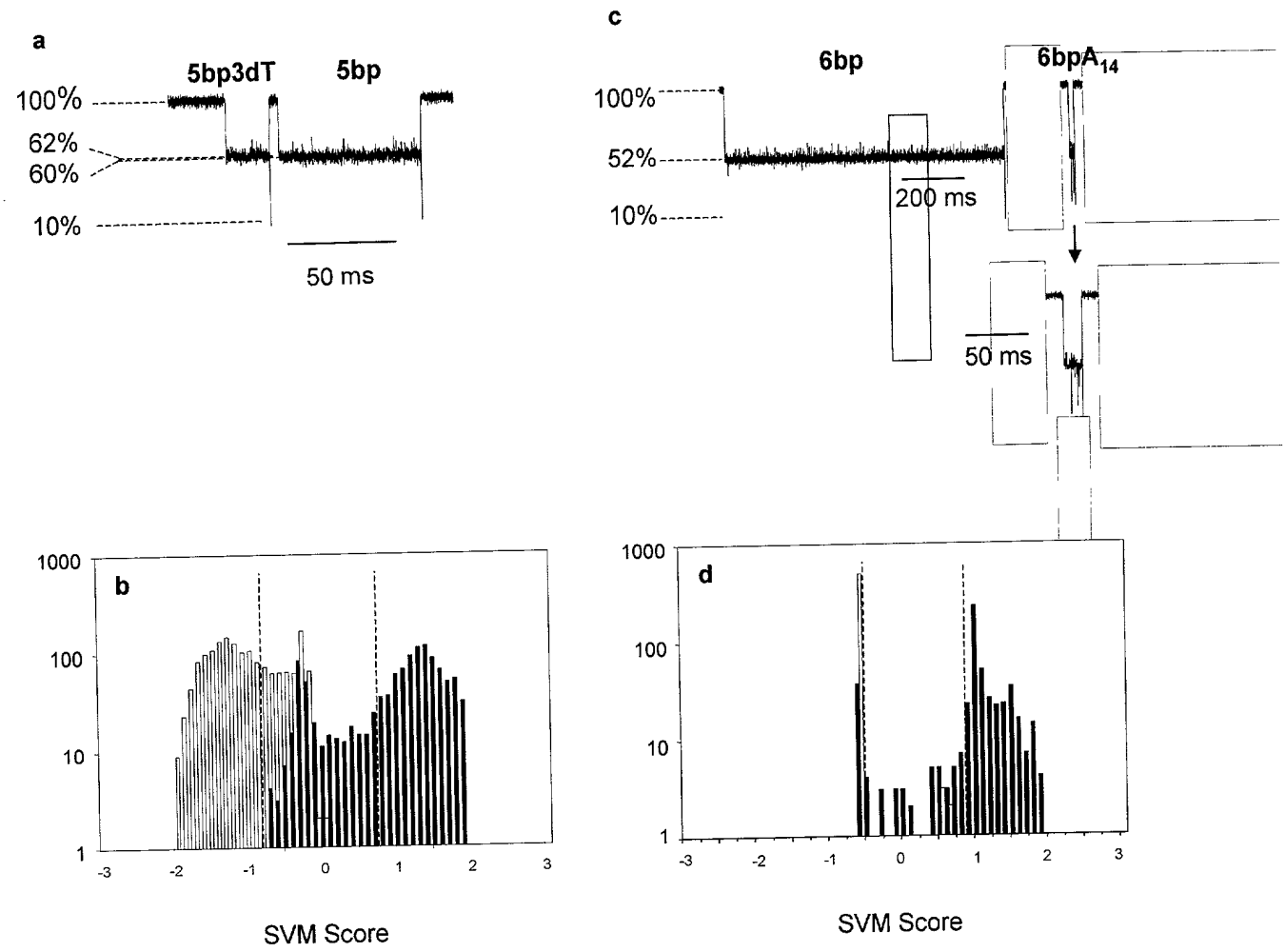

T T
T T
G:C
C:G
A:T
A:T
G:C
C:G
T:A
T:A
C:G
5' 3'
T T
T T
G:C
C:G
A:T
A:T
G:C
C:G
T:A
T:A
T:A
5' 3'
T T
T T
G:C
C:G
A:T
A:T
G:C
C:G
T:A
T:A
T:G
5' 3'
Current (pA)
160
120
80
40
0
-40
100
200
300
Time (ms)
0
100
200
300
Time (ms)
0
100
200
300
Time (ms)

Figure 3c



Number of Events

Figure 4

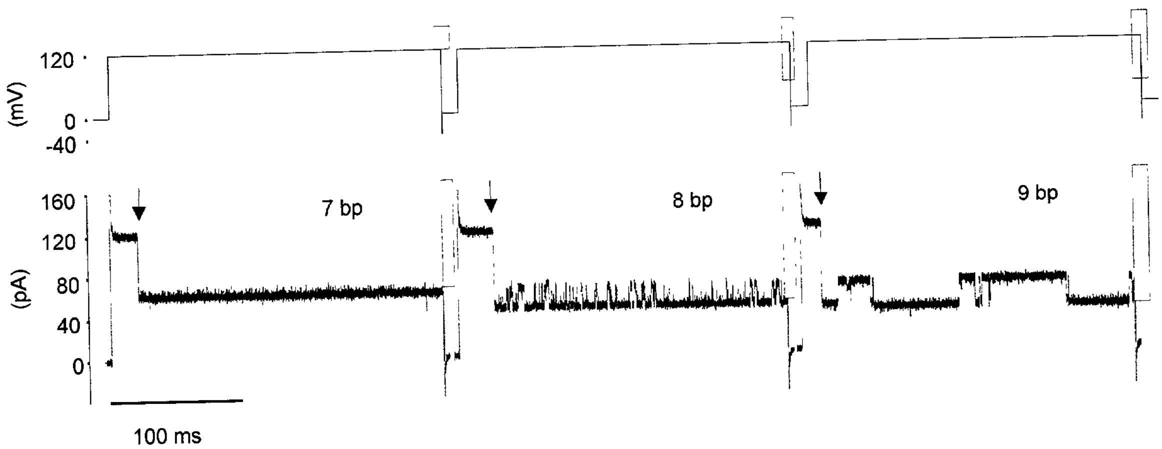
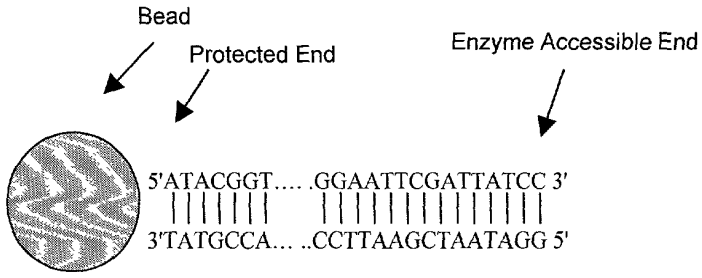


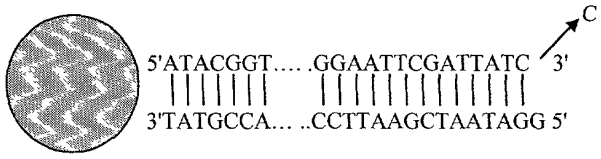
Figure 5

Figure 6.

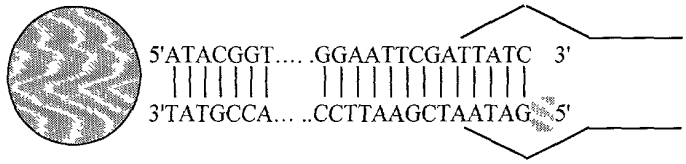
A) Blunt-ended DNAis attached at one end to a bead.



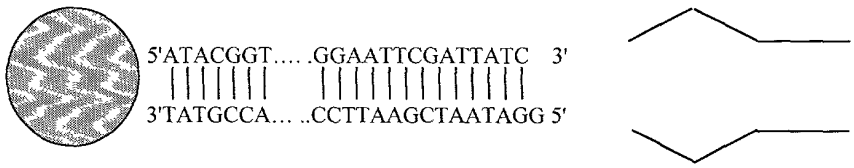
B) A single nucleotide is cut from the 3 end by a low processivity exonuclease such as exonuclease III.



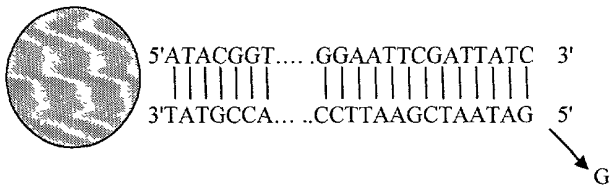
C) The single nucleotide overhang at the 5' end is read when the duplex end is captured in the nanoscale pore under an applied voltage.



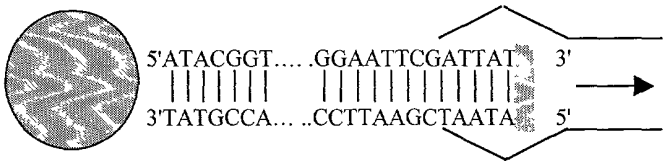
D)Once read, the DNA duplex is released from the nanopore by reversing the applied voltage.



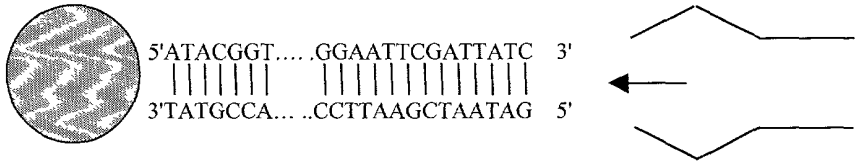
E) The single-nucleotide overhang is then cut with a nuclease (such as mung bean exonuclease), resulting in a blunt end.



F) The blunt end is then captured and held in the nanopore by an applied voltage. The terminal base-pair is identified while the duplex is captured.



G)Once read, the DNA duplex is released from the nanopore by reversing the applied voltage. The cycle is then repeated at step B).



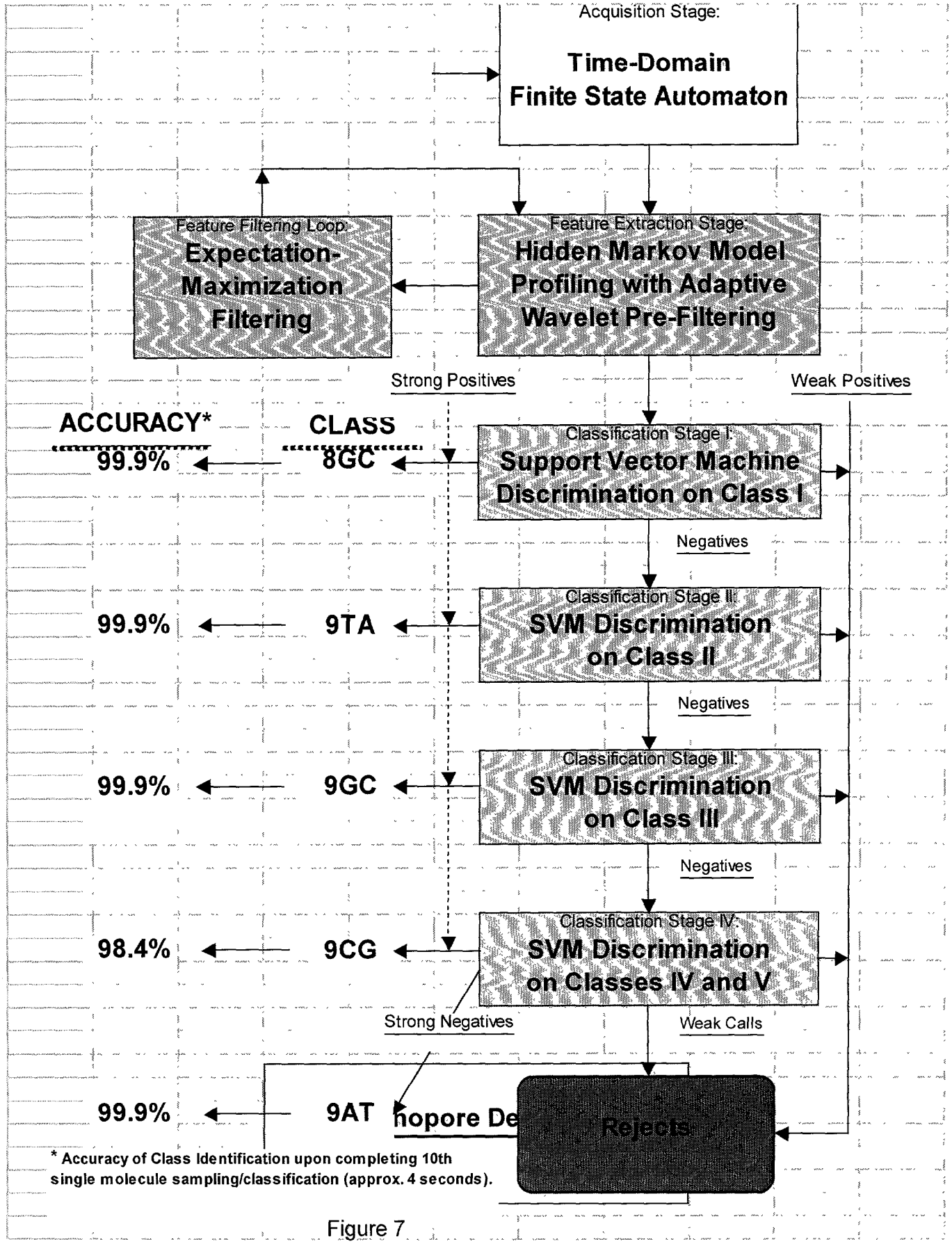
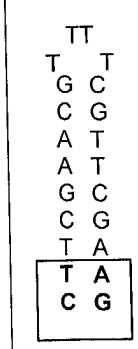
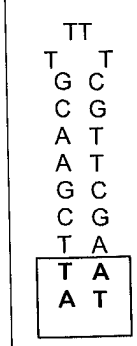
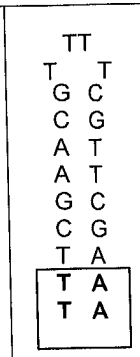
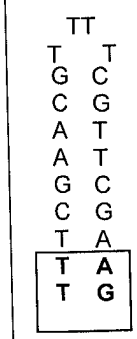
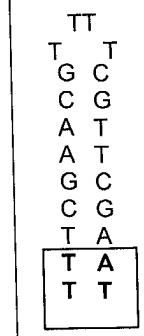
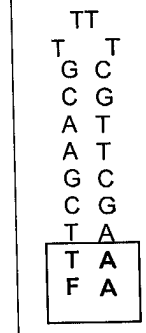
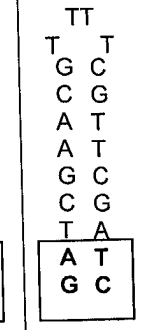
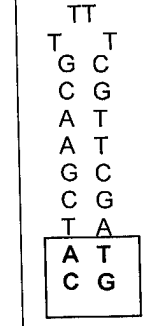
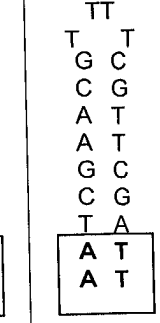
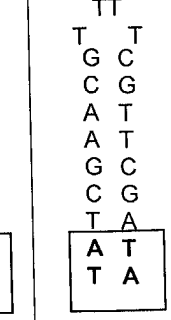

Acquisition Stage:
Time-Domain Finite State Automaton
Feature Filtering Loop:
Expectation-Maximization Filtering
Feature Extraction Stage:
Hidden Markov Model Profiling with Adaptive Wavelet Pre-Filtering
Strong Positives
Weak Positives
ACCURACY*
CLASS
99.9%
8GC
Classification Stage I:
Support Vector Machine Discrimination on Class I
Negatives
99.9%
9TA
Classification Stage II:
SVM Discrimination on Class II
Negatives
99.9%
9GC
Classification Stage III:
SVM Discrimination on Class III
Negatives
98.4%
9CG
Classification Stage IV:
SVM Discrimination on Classes IV and V
Strong Negatives
Weak Calls
99.9%
9AT
nopore De
Rejects
* Accuracy of Class Identification upon completing 10th single molecule sampling/classification (approx. 4 seconds).

Figure 7

Figure 8

Table 1. DNA hairpins used in this study. Primary sequence reads from 5′ end at bottom left to 3′ end at bottom right. Each hairpin has a 9 base-pair-long stem, and a four dT loop. The terminal base-pair and its nearest neighbor are highlighted by a box. These are the base-pairs in closest proximity to the pore limiting aperture when a given hairpin is captured in the α–hemolysin vestibule.

										
9bpGT/CA	9bpCT/GA	9bpAT/TA	9bpTT/AA	9bpTT/GA	9bpTT/TA	9bpFT/AA	9bpGA/CT	9bpCA/GT	9bpAA/TT	9bpTA/AT

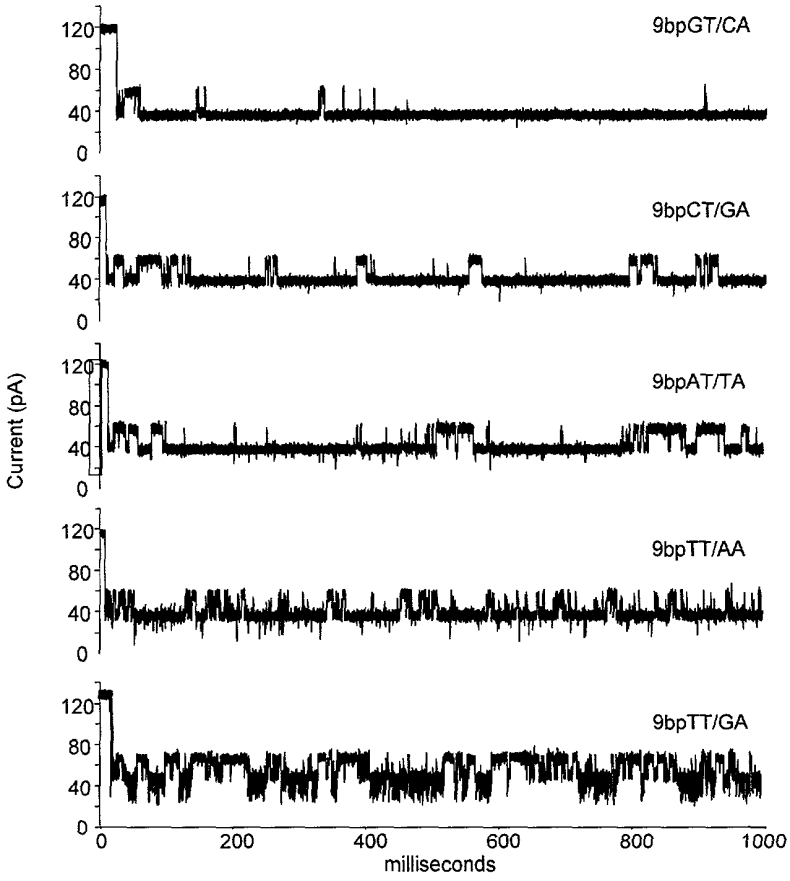

9bpGT/CA
9bpCT/GA
9bpAT/TA
9bpTT/AA
9bpTT/GA
120
80
40
0
Current (pA)
0
200
400
600
800
1000
milliseconds

Figure 9

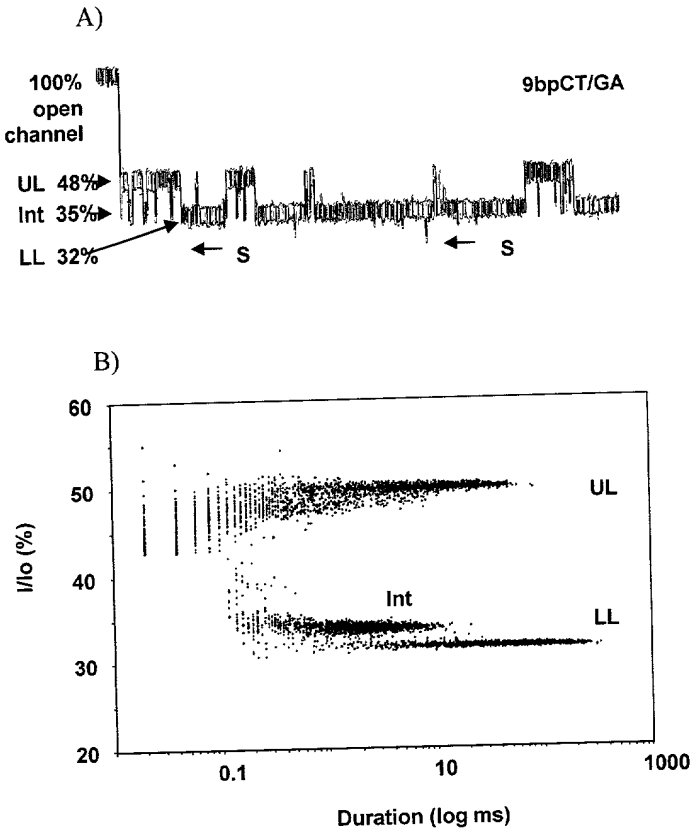

A)
100%
open
channel
9bpCT/GA
UL 48%
Int 35%
LL 32%
S
S
B)
60
50
40
30
20
I/Io (%)
UL
Int
LL
0.1
10
1000
Duration (log ms)

Figure 10

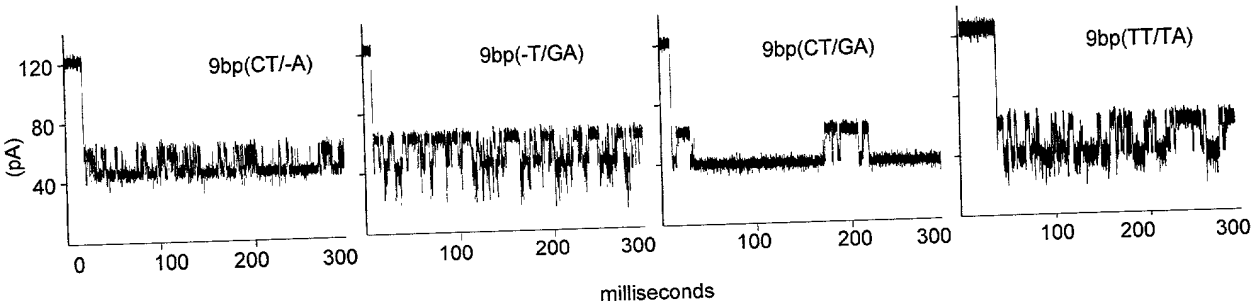

9bp(CT/-A)
9bp(-T/GA)
9bp(CT/GA)
9bp(TT/TA)
120
80
40
(pA)
0
100
200
300
milliseconds

Figure 11

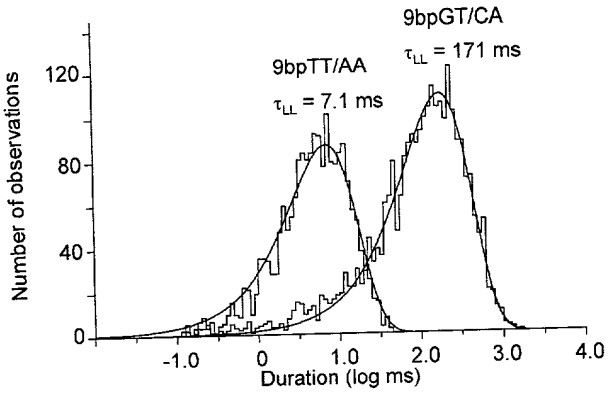

9bpGT/CA
τ_{LL} = 171 ms
9bpTT/AA
τ_{LL} = 7.1 ms
Number of observations
120
80
40
0
-1.0
0
1.0
2.0
3.0
4.0
Duration (log ms)

Figure 12

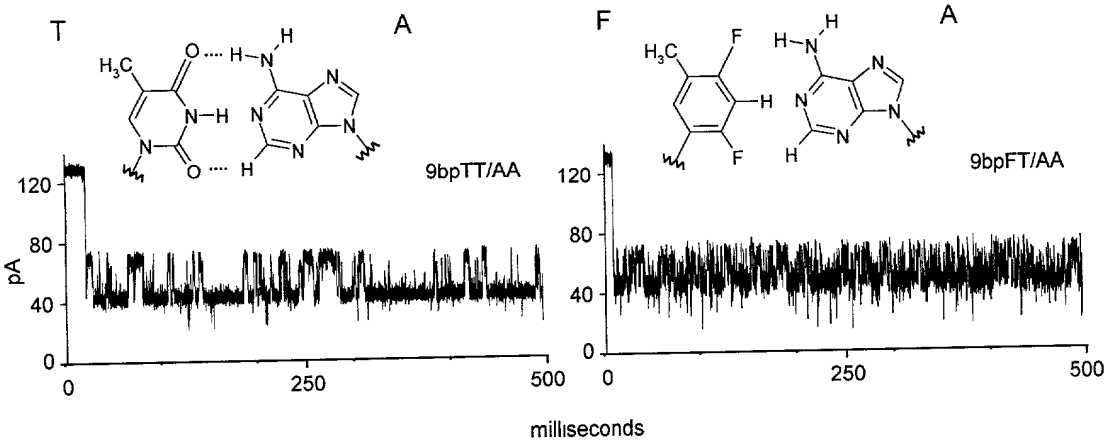


Figure 13

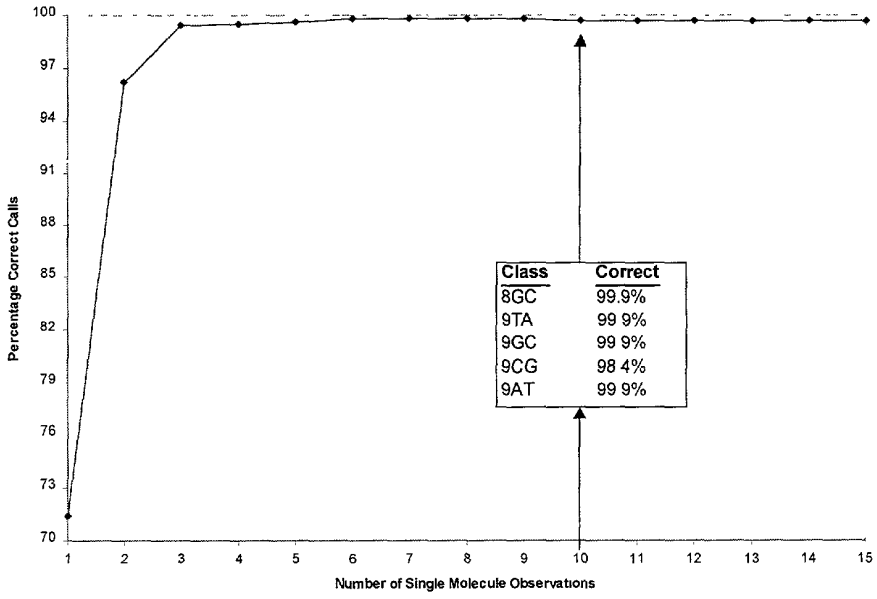

Percentage Correct Calls
100
97
94
91
88
85
82
79
76
73
70
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Number of Single Molecule Observations
Class
Correct
8GC
99.9%
9TA
99 9%
9GC
99 9%
9CG
98 4%
9AT
99 9%

Figure 14

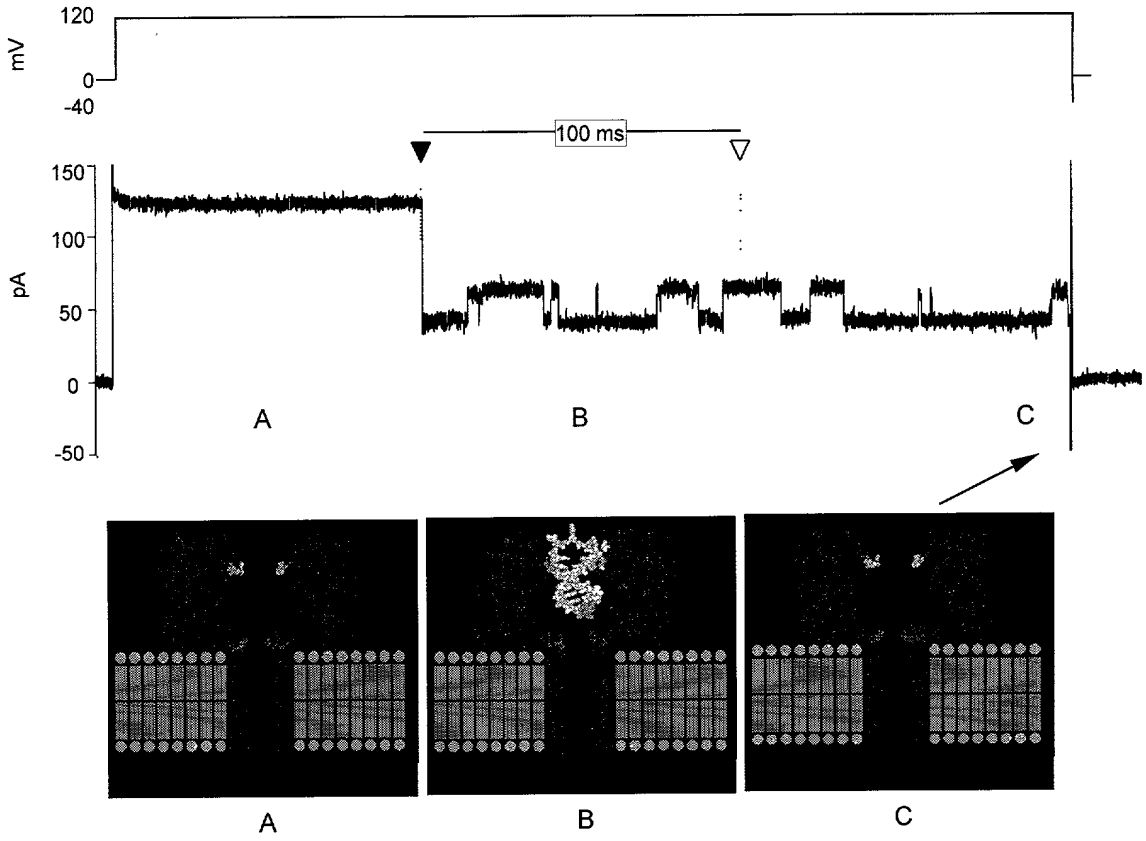


Figure 15.